

1                   **Perceptual learning of pitch direction in congenital amusia:**  
2                   **Evidence from Chinese speakers**

3  
4  
5                   Fang Liu<sup>a</sup>, Cunmei Jiang<sup>b,\*</sup>, Tom Francart<sup>c</sup>, Alice H. D. Chan<sup>d</sup>  
6                   and Patrick C. M. Wong<sup>e,f\*</sup>

7  
8                   <sup>a</sup>School of Psychology and Clinical Language Sciences, University of Reading,  
9                   Earley Gate, Reading RG6 6AL, UK

10                  <sup>b</sup>Music College, Shanghai Normal University, Shanghai, 200234, China

11                  <sup>c</sup>ExpORL, Department of Neurosciences, KU Leuven, B-3000 Leuven, Belgium

12                  <sup>d</sup>Division of Linguistics and Multilingual Studies, School of Humanities and Social  
13                  Sciences, Nanyang Technological University, S637332, Singapore

14                  <sup>e</sup>Department of Linguistics and Modern Languages and Brain and Mind Institute, The  
15                  Chinese University of Hong Kong, Shatin, N.T., Hong Kong SAR, China

16                  <sup>f</sup>The Chinese University of Hong Kong – Utrecht University Joint Center for  
17                  Language, Mind and Brain, Shatin, N.T., Hong Kong SAR, China

18  
19                  **Running head: Perceptual learning of pitch direction**

## **Abstract**

Congenital amusia is a lifelong disorder of musical processing for which no effective treatments have been found. The present study aimed to treat amusics' impairments in pitch direction identification through auditory training. Prior to training, twenty Chinese-speaking amusics and 20 matched controls were tested on the Montreal Battery of Evaluation of Amusia (MBEA) and two psychophysical pitch threshold tasks for identification of pitch direction in speech and music. Subsequently, ten of the twenty amusics undertook 10 sessions of adaptive-tracking pitch direction training, while the remaining 10 received no training. Post training, all amusics were re-tested on the pitch threshold tasks and on the three pitch-based MBEA subtests. Compared with those untrained, trained amusics demonstrated significantly improved thresholds for pitch direction identification in both speech and music, to the level of non-amusic control participants, although no significant difference was observed between trained and untrained amusics in the MBEA subtests. This provides the first clear positive evidence for improvement in pitch direction processing through auditory training in amusia. Further training studies are required to target different deficit areas in congenital amusia, so as to reveal which aspects of improvement will be most beneficial to the normal functioning of musical processing.

**Keywords:** congenital amusia; auditory training; pitch threshold; pitch direction; musical processing

## 1. Introduction

The ability to perceive music seems effortless and starts from infancy for the majority of the general population (Trehub, 2010). However, this ability can be beyond the reach of those with congenital amusia (amusia hereafter), a neurodevelopmental disorder of musical perception and production (Peretz, 2013). Often viewed as a lifelong disorder, individuals with amusia (amusics hereafter) demonstrate severe impairments in basic aspects of musical processing, such as distinguishing one tune from another and singing in tune, despite having normal hearing and intelligence and without any neurological or psychiatric disorders (Ayotte, Peretz, & Hyde, 2002). With a genetic origin (Drayna, Manichaikul, de Lange, Snieder, & Spector, 2001; Peretz, Cummings, & Dubé, 2007), this disorder affects around 1.5-5% of the general population for speakers of both tone and non-tonal languages (Kalmus & Fry, 1980; Nan, Sun, & Peretz, 2010; Peretz, 2013; Wong et al., 2012; but see Henry & McAuley, 2010, 2013 for criticisms). The core deficit of amusia lies in musical pitch processing, although around half of amusics also demonstrate rhythm deficits (Foxton, Nandy, & Griffiths, 2006; Hyde & Peretz, 2004; Peretz, Champod, & Hyde, 2003).

A range of perceptual skills are required for normal melodic processing, including acoustic analysis of pitch, extraction of interval and contour, “tonal encoding of pitch”, and short-term memory for pitch (Krumhansl & Keil, 1982; Peretz & Coltheart, 2003; Stewart, 2011). Amusics have shown impairments in all these aspects. First, amusics demonstrate difficulty in detecting pitch changes less than two semitones in tone sequences (Hyde & Peretz, 2004; Jiang, Hamm, Lim, Kirk, & Yang, 2011; Peretz et al., 2002), and show higher thresholds for pitch change detection than normal controls in psychophysical tasks (Foxton, Dean, Gee, Peretz, &

66 Griffiths, 2004; Jiang, Lim, Wang, & Hamm, 2013; Liu, Patel, Fourcin, & Stewart,  
67 2010). Second, amusics have reduced sensitivity to the direction of pitch movement  
68 (up versus down) in both music and speech, and show elevated psychophysical  
69 thresholds for pitch direction discrimination and identification (Foxton, Dean, et al.,  
70 2004; Jiang, Hamm, Lim, Kirk, & Yang, 2010; Jiang et al., 2013; Liu et al., 2010;  
71 Liu, Xu, Patel, Francart, & Jiang, 2012; Loui, Guenther, Mathys, & Schlaug, 2008).  
72 Third, amusics cannot detect out-of-key notes in Western music, or judge  
73 dissonance/consonance of musical excerpts (Ayotte et al., 2002; Peretz, Brattico,  
74 Järvenpää, & Tervaniemi, 2009). They are also impaired in explicit judgments of  
75 melodic expectation, musical syntax, and tonality relative to controls (Jiang, Liu, &  
76 Thompson, 2016; Omigie, Pearce, & Stewart, 2012; Zendel, Lagrois, Robitaille, &  
77 Peretz, 2015), despite demonstrating implicit processing of harmonic structure in  
78 priming tasks (Tillmann, Gosselin, Bigand, & Peretz, 2012). Finally, amusics show  
79 impaired short-term memory for pitch (Albouy, Mattout, et al., 2013; Tillmann,  
80 Schulze, & Foxton, 2009; Williamson & Stewart, 2010), which may result from their  
81 deficits in fine-grained pitch processing (Jiang et al., 2013).

82         A variety of theories have been put forward to explain the core deficits of  
83 amusia. One theory of amusia is that it is a disorder of top-down connectivity (Peretz,  
84 2013). This can be traced to disordered structure/function in the right inferior frontal  
85 gyrus (Hyde et al., 2007; Hyde, Zatorre, & Peretz, 2011), and disordered backwards  
86 connectivity from the inferior frontal gyrus to the auditory cortex (Albouy, Mattout, et  
87 al., 2013). Another theory, the “melodic contour deafness hypothesis” (Patel, 2008),  
88 proposes that reduced melodic contour (or pitch direction) perception in amusia may  
89 have prevented amusics from learning musical intervals and perceiving melodic  
90 structure.

91 Previous evidence indicates that the amusic brain only has “limited plasticity”  
92 in response to music training/listening (Peretz, 2013). Several single case reports  
93 documented null results of regular music/piano lessons and singing in choirs and  
94 school bands on amusia (Allen, 1878; Geschwind, 1984; Lebrun, Moreau, McNally-  
95 Gagnon, Mignault Goulet, & Peretz, 2012; Peretz et al., 2002). Two recent studies  
96 also examined the effects of daily music listening (Mignault Goulet, Moreau,  
97 Robitaille, & Peretz, 2012) and weekly singing intervention (Anderson, Himonides,  
98 Wise, Welch, & Stewart, 2012) on musical processing in amusia, with the numbers of  
99 amusic participants being 8 (Mignault Goulet et al., 2012) and 5 (Anderson et al.,  
100 2012), respectively. Neither study included an untrained amusic group as a control  
101 group. In (Mignault Goulet et al., 2012), after four weeks of daily half-hour listening  
102 of popular songs, the eight 10-13 year old amusic children showed no improvement in  
103 either behavioral (pitch change detection) or neural (the P300 component) measures  
104 of pitch processing. Thus, daily music listening does not seem to be an effective  
105 strategy to reduce amusic symptoms (Mignault Goulet et al., 2012). Similarly, after  
106 seven weekly group-singing workshops, which incorporated learning activities such  
107 as vocal warm-ups and listening of melodies on pianos/keyboards combined with  
108 three or four 15-min sessions of self-exercises with *Sing and See* per week at home,  
109 the five amusics in (Anderson et al., 2012) only improved in singing of the familiar  
110 song “*Happy birthday*”, but not in any other measures such as computer and vocal  
111 pitch matching, MBEA scale subtest, or singing of the self-chosen song. Together,  
112 these results suggest that passive exposure to musical stimuli and general-purpose  
113 singing or music training methods are not appropriate remediation strategies for  
114 individuals with congenital amusia, who have impoverished auditory and memory  
115 resources, at least not at the dosage that was prescribed.

However, the fact that humans can improve perception skills through learning and practice is well documented across all sensory modalities, including auditory (Wright & Zhang, 2009), visual (Gilbert & Li, 2012), tactile (M. Wong, Peters, & Goldreich, 2013), olfactory (Gottfried, 2008), and taste (Peron & Allen, 1988). Music training, in particular, has been shown to enhance both musical and speech processing, and induce substantial neurophysiological, neuroanatomical, and functional changes in the human brain across the lifespan (Herholz & Zatorre, 2012; Patel, 2011). It is thus surprising that the amusic brain would be less malleable than neurotypical brains in perceptual learning.

Several factors might be responsible for the “limited plasticity” of the amusic brain documented in past research. First, the music training/listening activities reported in previous studies did not tap directly into individual target deficit areas of amusia, e.g., impaired fine-grained pitch discrimination, insensitivity to pitch direction, and lack of pitch awareness (Loui et al., 2008; Loui, Kroog, Zuk, & Schlaug, 2011; Patel, 2008; Peretz et al., 2002, 2009; Stewart, 2008), but instead employed general-purpose music training methods such as daily music listening (Mignault Goulet et al., 2012), singing in choirs or school bands (Lebrun et al., 2012; Peretz et al., 2002), taking regular music/piano lessons (Allen, 1878; Geschwind, 1984), or using a broad-brush singing intervention approach (Anderson et al., 2012). These methods, although useful, may take months or years to make significant effects (Besson, Schön, Moreno, Santos, & Magne, 2007; Herholz & Zatorre, 2012; Patel, 2011), especially for amusics who have widespread musical disorders. On the other hand, in the field of language acquisition, it has been found that successful learning benefits from starting small (Elman, 1993; Goldowsky & Newport, 1993). That is, young children, with limited cognitive and memorial capabilities, may learn language

through analyzing the components of complex stimuli, rather than performing a holistic analysis of the whole form like adults do (Newport, 1988). Given the limited auditory and memory capacities for musical processing in amusia, it is possible that the amusic brain is too overwhelmed to benefit from the vast amount of information embedded in those general-purpose music training/listening activities. Alternative approaches targeting core deficit areas of amusia might be able to help treat amusia.

Pitch direction is a building block of melodic contour (Patel, 2008; Stewart, 2008), which is in turn one of the most important features for the perception and storage of melody in memory (Dowling, 1978; Dowling & Fujitani, 1971; Idson & Massaro, 1978). Based on the hypothesis that amusia is at least partially due to insensitivity to the direction of pitch movement (Loui et al., 2008; Stewart, 2008), or the “melodic contour deafness hypothesis” (Patel, 2008), it is likely that the pitch direction deficit in amusia has led to developmental problems with perception of melodic contour and music as a whole (Patel, 2008).

To assess the processing of pitch direction in amusia, we have used two different types of tasks in our previous studies: pitch direction *discrimination* (Liu et al., 2010 on English speakers; Liu, Jiang, et al., 2012 on Mandarin speakers), and pitch direction *identification* (Liu, Xu, et al., 2012). In the *discrimination* task (Liu, Jiang, et al., 2012; Liu et al., 2010), participants were asked to report which of the three gliding tones differed in direction from the other two (e.g., the “falling” glide in the “rising-rising-falling” sequence, AXB task), thus *discriminating* the direction of pitch change. Furthermore, in the *discrimination* task, labelling of tone patterns as rising or falling was not required, and participants were simply requested to report which was the “odd one out” in pitch direction in a sequence of three tones. In the *identification* task (Liu, Xu, et al., 2012), only two tones were presented in one trial,

and participants were required to *identify* the direction of pitch movement (e.g., high-low versus low-high, two-alternative forced-choice task). For pitch direction *discrimination* (Liu, Jiang, et al., 2012; Liu et al., 2010), both Mandarin-speaking amusics and controls achieved lower (better) pitch thresholds than their English-speaking counterparts. This superior performance on pitch direction *discrimination* in Mandarin speakers may result from passive perceptual learning of this sound feature in their native language (Liu, Jiang, et al., 2012). However, for pitch direction *identification* (Liu, Xu, et al., 2012), which requires conscious pitch direction awareness, both Mandarin-speaking amusics and controls showed elevated thresholds compared to pitch direction *discrimination* (Liu, Jiang, et al., 2012). This suggests that pitch direction *identification* is a more difficult (or cognitively demanding) task than pitch direction *discrimination*, even for tone language speakers, and especially for amusics.

Aiming to enhance amusics' fine-grained pitch discrimination, pitch direction recognition, and pitch awareness, we designed and implemented an auditory training program to help amusics recognize pitch direction in music and speech. We hypothesized that training and improvement on pitch direction identification would provide the scaffolding for amusics to build complex musical systems, and consequently help ameliorate musical processing deficits in amusia.

## **2. Materials and Methods**

### **2.1. Participants**

Twenty Chinese-speaking amusics and 20 control participants were recruited through advertisements posted on the university bulletin board systems and mass mail services in Shanghai and Hong Kong, China. The Montreal Battery of Evaluation of Amusia (MBEA) (Peretz et al., 2003) was used to diagnose amusia in these



participants. Consisting of six subtests, the MBEA measures the perception of scale, contour, interval, rhythm, meter, and memory of melodies. Participants were classified as amusic if scored 65 or under on the pitch composite score (sum of the scores on the scale, contour, and interval subtests) or below 78% correct on the MBEA global score, which corresponds to 2 standard deviations below the mean score of normal controls (Liu et al., 2010; Peretz et al., 2003). Participants in the control group were chosen to match with the amusic group in sex, handedness, age, music training background, and years of education, but having MBEA scores within the normal range. Before conducting the experiments, the amusic group was randomly divided into two subgroups: trained amusics ( $n = 10$ ) were asked to participate in our pitch direction training program, whereas untrained amusics ( $n = 10$ ) received no training. Table 1 summarizes the characteristics of the amusic (trained versus untrained) and control groups. As can be seen, controls performed significantly better than amusics on the MBEA. Although trained amusics received more years of education than the untrained ( $p = .01$ ), the two groups did not differ significantly in the MBEA at the pretest. Years of education was used as a covariate in the linear mixed-effects models as described in the Results section. None of the participants reported having speech or hearing disorders or neurological/psychiatric impairments in the questionnaires concerning their music, language, and medical background. All were undergraduate or postgraduate students at universities in Shanghai or Hong Kong, with Mandarin Chinese or Cantonese as their native language, and none had received any formal extracurricular music training. Ethical approvals were granted by Shanghai Normal University and The Chinese University of Hong Kong. Written informed consents were obtained from all participants prior to the experiment.

[Insert Table 1 about here]

## 2.2. Tasks

The experiment consisted of a practice session (with audiovisual feedback), a pre-training test (pretest hereafter; with no feedback), 10 training sessions (with audiovisual feedback), and a post-training test (posttest hereafter; with no feedback). Tasks involved identification of pitch direction (high-low versus low-high) in pairs of sounds with varying pitch distances using two-interval forced-choice (2IFC) methods, with procedure adapted from our previous study (Liu, Xu, et al., 2012).

In particular, in the current study, we modified the protocol in Liu, Xu, et al. (2012) by using the “two-down one-up” staircase method (instead of “three-down one-up” in Liu, Xu, et al., 2012) and piano tones (instead of complex tones in Liu, Xu, et al., 2012). We also excluded gliding pitches (e.g., rising-falling, falling-rising), as amusics had less difficulty recognizing pitch direction in gliding than in discrete pitches, for both speech and non-speech stimuli (Liu, Xu, et al., 2012). Fig. 1 shows the schematic diagram of stimulus presentation, with each stimulus lasting 250 ms separated by an inter-stimulus interval of 250 ms. Participants were instructed to choose between two choices given on the computer screen (via mouse click) to indicate the pitch pattern of the stimulus pair: “高低” (“high low”) or “低高” (“low high”).

[Insert Fig. 1 about here]

Control participants ( $n = 20$ ) were administered the practice session and pretest only. All amusics ( $n = 20$ ) completed the practice session, pretest, and posttest (pre- and post-test were about two weeks apart). The two amusic groups were comparable in pitch thresholds at pretest: thresholds for speech syllable:  $t(18) = -0.74$ ,  $p = .47$ ; thresholds for piano tone:  $t(18) = 0.57$ ,  $p = .58$ . In order to see whether training in pitch direction identification would improve musical pitch processing, all

amusics (trained or untrained) were also re-tested on the first three subtests (scale, contour, and interval) of the MBEA.

### **2.3. Stimuli**

Stimuli were of two types, the Mandarin/Cantonese syllable /ma/ and its piano tone analog. Our stimuli were based on sounds with level pitches, since these occur both in music and in the level tones of Mandarin and Cantonese (Duanmu, 2007; Yip, 2002). It has been shown that Mandarin speakers with amusia have difficulty in identifying/discriminating lexical tones and pitch direction in speech and music (Liu, Jiang, et al., 2012; Liu, Xu, et al., 2012; Nan et al., 2010). We thus used two different stimulus types to ensure that pitch direction training was done for both domains.

For each stimulus type, one single token was used to create all stimuli with different pitches. The original speech syllable /ma/ was produced by a male native speaker of Mandarin (Liu, Xu, et al., 2012), and its piano tone analog was generated using a Virtual Grand Piano, Pianissimo (Acoustica, Inc.). The durations of the two original stimuli were then normalized to 250 ms, and their fundamental frequencies were manipulated to include a range of pitches from 131 Hz (corresponding to the note C3 on the musical scale) to 330 Hz (note E4) using a custom-written script for the Praat program (Boersma & Weenink, 2001). Since the effect of intensity on tone perception is negligible when pitch is present (Lin, 1988) and in keeping with previous studies on speech/pitch processing in amusia (Ayotte et al., 2002; Jiang et al., 2010; Liu et al., 2010; Liu, Xu, et al., 2012; Loui et al., 2008; Patel, Foxton, & Griffiths, 2005; Patel, Wong, Foxton, Lochy, & Peretz, 2008), we intentionally did not manipulate the amplitude of the stimuli in order to preserve the natural quality of these sounds.

For both stimulus types, there were a standard stimulus of 131 Hz (C3) and 63

target stimuli that deviated from the standard in steps ( $\Delta F$ ,  $F_0$  difference or pitch interval between the standard and target stimuli) of 0.01 (10 steps between 131.08 and 131.76 Hz, increasing by 0.01 semitones in each step), 0.1 (9 steps between 131.76 and 138.79 Hz, increasing by 0.1 semitones in each step), and 0.25 semitones (44 steps between 138.79 and 262 Hz, increasing by 0.25 semitones in each step). Thus, the smallest pitch interval ( $\Delta F$  between the standard and step 1 deviant) between the standard and target stimuli was 0.01 semitones, and the largest pitch interval ( $\Delta F$  between the standard and step 63 deviant) was 12 semitones in the testing/training sessions.

#### **2.4. Procedure**

The practice sessions (for both speech syllable and piano tone) consisted of 8 trials, with pitch intervals (13-16 semitones) greater than those in the testing/training sessions. The trials were presented in a random order with no adaptive tracking procedure applied. Participants were required to achieve 100% correct on the practice trials (with audiovisual feedback) before proceeding to the testing sessions.

In both testing and training sessions, stimuli were presented with adaptive tracking procedures using the APEX 3 program developed at ExpORL (Francart, van Wieringen, & Wouters, 2008). As a test platform for auditory psychophysical experiments, APEX 3 enables the user to specify custom stimuli and procedures with eXtensible Markup Language (XML). The “two-down, one-up” staircase method was used in the adaptive tracking procedure, with step sizes of 0.01, 0.1, and 0.25 semitones as explained earlier. Following a response, the next trial was played 750 ms later. In the staircase, a reversal was defined when there was a change of direction, e.g., from “down” to “up”, or from “up” to “down”. Each run would end after 14 such reversals, and the threshold (in semitones) was calculated as the mean of the pitch

291 intervals (pitch differences between the standard and target stimuli) in the last 6  
292 reversals. Across all participants, it took on average 6.67 minutes (SD = 2.03) and  
293 6.35 minutes (SD = 1.29) to complete pre- and post-tests for piano tone thresholds,  
294 and 7.51 minutes (SD = 8.00) and 6.83 minutes (SD = 2.58) for speech syllable  
295 thresholds.

296 As mentioned earlier, ten of the twenty amusics were assigned to the training  
297 group, and completed 10 training sessions of pitch direction identification over  
298 around two weeks. These training sessions were administered on different days, with  
299 no more than two days between consecutive sessions. Each session lasted about 30  
300 minutes. The starting pitch interval ( $\Delta F$ ) between the standard and target stimuli was  
301 12 semitones for the first two training sessions, which consisted of one run of each  
302 stimulus type (speech syllable and piano tone). Starting from the third training  
303 session, an adaptive training protocol was used, in which the participant's threshold  
304 on an earlier run (the average step of the last 6 reversals) was taken as the initial step  
305 for the next run. This adaptive training protocol ensured that trained pitch intervals  
306 were adjusted based on participants' performance over time. Given the increased  
307 difficulty (near-threshold) of the trained pitch intervals during adaptive training, it  
308 took less time for the 14 reversals in each run to complete, and thus the duration of  
309 each run became much shorter. Consequently, two runs of speech syllable and piano  
310 tone were administered in training sessions 3-10, compared to one run each in training  
311 sessions 1-2.

312 Participants received feedback during training. The text "Correct. :)" was  
313 displayed following correct responses, and "Incorrect. :( " was shown for incorrect  
314 responses. In either case, the correct answer ("低高" or "高低", "low-high" or "high-  
315 low") together with its graphic representation was shown to the participants on the

computer screen. After seeing the feedback, participants could choose to play the trial again, or go directly to the next trial.

All stimuli were presented diotically via Philips SHM1900 headphones (in Shanghai) or Sennheiser HD 380 PRO Headphones (in Hong Kong) at a comfortable listening level. The order of speech syllable and piano tone blocks was counterbalanced across participants and runs/sessions.

## 2.5. *Statistical analyses*

Statistical analyses were conducted using R (R Core Team, 2014). Thresholds were transformed using log transformation for parametric statistical analysis (Howell, 2009), as amusics' thresholds deviated significantly from normal distributions (Shapiro-Wilk normality test: pretest for piano tones:  $W = 0.86$ ,  $p = .008$ ; pretest for speech syllables:  $W = 0.73$ ,  $p < .001$ ; posttest for piano tones:  $W = 0.67$ ,  $p < .001$ ; posttest for speech syllables:  $W = 0.63$ ,  $p < .001$ ). In order to account for the possible contribution of education to the current results (the two amusic subgroups differed in years of education as shown in Table 1), years of education were entered as a covariate in the linear mixed-effects models in the Results section. Although there was also a difference in age between the two groups ( $p = .06$ , Table 1), age was not included in the mixed-effects models due to the collinearity between age and education in the amusic participants ( $r(18) = .79$ ,  $p < .001$ ). Effect sizes in the ANOVA models were calculated using generalized eta squared,  $\eta_G^2$  (Bakeman, 2005; Olejnik & Algina, 2003), and those in  $t$ -tests were calculated using Cohen's  $d$  (Cohen, 1988). Following (Cohen, 1988), an  $\eta_G^2$  above .02 ( $d > 0.20$ ) reflects a small effect, an  $\eta_G^2$  above .13 ( $d > 0.50$ ) reflects a medium effect, and an  $\eta_G^2$  above .26 ( $d > 0.80$ ) reflects a large effect (Bakeman, 2005). Post-hoc pairwise comparisons were conducted using two-tailed  $t$  tests with  $p$ -values adjusted using the Holm method

(Holm, 1979).

### 3. Results

Fig. 2 shows mean pitch direction identification thresholds of amusics and controls at pre- and post-tests for piano tones and speech syllables. A linear mixed-effects model was conducted on log-transformed thresholds of the two amusic groups, with training (trained versus untrained) as the between-subjects factor, education as a covariate, stimulus type (speech syllable versus piano tone) and test (pretest versus posttest) as within-subjects factors, and participants (trained and untrained amusics) as random effects (see Supplementary Table 1 for detailed results). Results revealed significant effects of test ( $F(1,48) = 30.42, p < .001$ ) and training ( $F(1,16) = 16.46, p < .001$ ), as posttest thresholds were significantly lower (better) than pretest thresholds and trained amusics achieved better thresholds than untrained amusics. The main effects of education ( $F(1,16) = 2.85, p = .11$ ) and stimulus type ( $F(1,48) = 2.21, p = .14$ ) were not significant. A significant test  $\times$  training interaction ( $F(1,48) = 18.50, p < .001$ ) was observed, owing to the fact that thresholds did not differ between trained and untrained amusics at pretest ( $p = .92$ ) but trained amusics showed significantly lower (better) thresholds than untrained amusics at posttest ( $p < .001$ ). There was also a significant stimulus type  $\times$  training interaction ( $F(1,48) = 7.17, p = .01$ ), as thresholds (pre- and post-test combined) did not differ between trained and untrained amusics for speech syllables ( $p = .33$ ), but the two groups differed significantly in thresholds for piano tones ( $p = .01$ ). Other interactions were not significant (all  $ps > .05$ ).

Two sample  $t$ -tests (two-sided) were conducted to see how the two amusic groups compared with controls in thresholds at pre- and post-test. At pretest, controls outperformed the two amusic groups for both piano tones (trained amusics vs.

controls:  $t(28) = 8.31, p < .001, d = 3.22$ ; untrained amusics vs. controls:  $t(28) = 6.02, p < .001, d = 2.33$ ) and speech syllables (trained amusics vs. controls:  $t(28) = 5.55, p < .001, d = 2.15$ ; untrained amusics vs. controls:  $t(28) = 6.03, p < .001, d = 2.34$ ). When amusics' posttest thresholds were compared with controls' pretest thresholds, untrained amusics showed worse performance than controls on both tasks (piano tones:  $t(28) = 4.99, p < .001, d = 1.93$ ; speech syllables:  $t(28) = 5.57, p < .001, d = 2.16$ ), whereas trained amusics achieved similar thresholds as controls (piano tones:  $t(28) = 1.61, p = .12, d = 0.62$ ; speech syllables:  $t(28) = -0.60, p = .55, d = 0.23$ ).

[Insert Fig. 2 about here]

Fig. 3 shows mean pitch thresholds across the 10 training sessions for the 10 trained amusics for piano tones and speech syllables. A repeated measures ANOVA suggested that amusic thresholds significantly improved over 10 training sessions [ $F(9,81) = 23.10, p < .001$  after correction using Greenhouse-Geisser epsilon,  $\eta_G^2 = .47$ ]. There was no significant effect of stimulus type [ $F(1,8) = 2.55, p = .15, \eta_G^2 = .02$ ] or stimulus type  $\times$  session interaction [ $F(9,81) = 0.33, p = .79$  after correction using Greenhouse-Geisser epsilon,  $\eta_G^2 = .01$ ]. This indicates that trained amusics improved on pitch direction identification thresholds for piano tones and speech syllables at similar rates over the 10 training sessions. Post-hoc analysis ( $p$ -values adjusted using the Holm method) indicated that trained amusics' thresholds differed significantly between sessions 1 and 2-10 (all  $ps < .01$ ), between sessions 2 and 1, 4-10 (all  $ps < .05$ ), and between sessions 3 and 1, 9 (both  $ps < .05$ ). Other pairwise comparisons were non-significant (all  $ps > .05$ ). This pattern of improvement may be due to the adaptive training protocol we used after training session 3: the starting pitch interval for sessions 3-10 was determined by the threshold obtained from the previous run, and each run always ended after 14 reversals. On the one hand, this



ensured that participants were trained on pitch intervals centered on their thresholds. On the other hand, this made the resultant thresholds in sessions 1-2 (the starting pitch interval was 12 semitones) and 3-10 (the starting pitch interval was at threshold) largely incomparable.

[Insert Fig. 3 about here]

In order to see the role of pretest threshold in predicting posttest threshold, a linear mixed-effects model was fit on posttest threshold with training (trained versus untrained) and stimulus type (piano tone versus speech syllable) as fixed effects, pretest threshold and education as covariates, and participants (trained and untrained amusics) as random effects (see Supplementary Table 2 for detailed results). Results revealed a significant effect of training ( $F(1,16) = 135.57, p < .001$ ), despite the fact that pretest threshold ( $F(1,8) = 54.80, p < .001$ ) and education ( $F(1,16) = 18.36, p < .001$ ) also strongly predicted posttest threshold. There was also a significant training  $\times$  pretest threshold interaction ( $F(1,8) = 26.87, p < .001$ ), as posttest thresholds of trained amusics were less affected by pretest thresholds than untrained amusics. This was confirmed by different correlations between pre- and post-test pitch thresholds for trained versus untrained amusics (Figure 4). For trained amusics, pre- and post-test thresholds did not correlate for either piano tones ( $r(8) = .52, p = .13$ ) or speech syllables ( $r(8) = .48, p = .16$ ), due to improvement from training. In contrast, untrained amusics showed significant positive correlations between pre- and post-test thresholds for both piano tones ( $r(8) = .66, p = .04$ ) and speech syllables ( $r(8) = .87, p = .001$ ), which suggests that untrained amusics tended to perform similarly at pre- and post-tests. Finally, there was a significant stimulus type  $\times$  training  $\times$  pretest threshold interaction ( $F(1,8) = 6.55, p = .03$ ), as trained amusics' post-test thresholds for speech syllables were less affected by pre-test thresholds than for piano tones.

Other effects/interactions were not significant.

[Insert Fig. 4 about here]

Fig. 5 plots mean scores of the 10 trained and 10 untrained amusics for MBEA scale, contour, and interval subtests at pre- and post-tests. These three subtests measure individuals' abilities to process scale structure, melodic contour, and pitch interval in Western melodies, respectively (Peretz et al., 2003). A linear mixed-effects model was fit on posttest MBEA score with training (trained versus untrained) and task (scale, contour, and interval) as fixed effects, pretest score and education as covariates, and participants (trained and untrained amusics) as random effects (see Supplementary Table 3 for detailed results). Results revealed a significant main effect of education ( $F(1,16) = 7.26, p = .02$ ), as posttest MBEA scores showed a negative correlation with years of education participants received ( $r(58) = -.23, p = .08$ ). There was also a significant interaction between education and pretest score ( $F(1,20) = 5.28, p = .03$ ), while other effects/interactions were not significant (all  $ps > .05$ ). Planned contrasts (with the directional hypothesis of training induced improvement) indicated that trained amusics significantly improved on the MBEA contour subtest ( $t(9) = 2.10, p = .03$ , one-tailed,  $d = 0.66$ ), but not on scale or interval subtests (both  $ps > .05, ds < 0.50$ ). No improvement was observed in untrained amusics on any of the three MBEA subtests (all  $ps > .10, ds < 0.50$ ). However, at posttest, trained and untrained amusics did not differ significantly for any of the three MBEA subtests (all  $ps > .05, ds < 0.50$ ). Correlation analyses revealed no significant correlations between pre- and post-test MBEA scale/contour/interval scores for either trained or untrained amusics (all  $ps > .10$ ). This was due to the random variations in pre- and post-test MBEA scores within and across participants (Figure 6).

[Insert Fig. 5 about here]

[Insert Fig. 6 about here]

In order to see whether controls' baseline performance on the pitch threshold tasks was optimized or not, we trained one control participant (C1) using the same protocol as used for the amusics. No improvement was observed from pre- to post-test for either piano tone (0.10 vs. 0.12 st) or speech syllable (0.14 vs. 0.15 st). Although we are unable to reach a definitive conclusion with only one participant, it appears that the accurate minimum thresholds for the current tasks should approximate the best controls' performance.

#### **4. Discussion**

Suffering from a lifelong disorder of musical perception and production, individuals with congenital amusia have only shown "limited plasticity" in response to music training/listening in past research (Peretz, 2013). Tapping into the core deficits of amusia and using a scaffolding, incremental learning approach, the present study investigated whether amusics' pitch direction identification thresholds could be improved, and if so, whether enhanced pitch direction recognition would facilitate musical processing in amusia. To this end, we designed an adaptive-tracking training paradigm to help amusics consciously label the direction of fine-grained pitch movement in both speech syllables and piano tones. After undertaking 10-session training programs over two weeks, trained amusics demonstrated significantly improved thresholds for pitch direction identification in both speech syllables and piano tones. However, although trained amusics demonstrated better performance on the contour subtest of the MBEA at posttest compared to pretest, no significant difference was observed between trained and untrained amusics in any of the three pitch-based MBEA subtests. These findings provide the first evidence for the improvement of pitch direction perception in amusia, although this may not lead to

improved musical processing. This not only opens possibilities for designing other rehabilitative programs to treat this musical disorder, but also has significant implications for theories and applications in music and speech learning.

Previous evidence indicates that the amusic brain only has “limited plasticity” in response to music training/listening (Peretz, 2013), be it singing training, regular music/piano lessons, daily musical listening, or being involved in choirs or school bands (Allen, 1878; Anderson et al., 2012; Geschwind, 1984; Lebrun et al., 2012; Mignault Goulet et al., 2012; Peretz et al., 2002). This may be due to the fact that, with limited auditory and memory capacities, individuals with congenital amusia are unable to benefit from passive exposure to musical stimuli or general-purpose singing or music training methods. In light of the “less is more hypothesis” in language acquisition (Elman, 1993; Goldowsky & Newport, 1993) and the pitch direction or “melodic contour deafness” hypothesis in amusia (Loui et al., 2008; Patel, 2008; Stewart, 2008), the current investigation used a scaffolding approach and conducted the first auditory training study to explore whether pitch direction identification could be improved through perceptual learning, and if yes, whether it could further help ameliorate musical processing deficits in amusia. After 10 sessions, trained amusics showed improved pitch direction identification thresholds, but did not outperform untrained amusics in musical processing, as indexed by the three pitch-based MBEA subtests. This suggests that improvement in pitch direction processing does not necessarily entail improvement in musical processing.

Thus, it is worth noting that the ability to discriminate pitch direction develops with age in children (Fancourt, Dick, & Stewart, 2013). Apart from amusics, some typical adult listeners also show difficulty in pitch direction recognition (Foxton, Weisz, Bauchet-Lecaignard, Delpuech, & Bertrand, 2009; Mathias, Bailey, Semal, &

Demany, 2011; Mathias, Michey, & Bailey, 2010; Neuhoff, Knight, & Wayand, 2002; Semal & Demany, 2006), so do individuals with developmental dyslexia (Ziegler, Pech-Georgel, George, & Foxton, 2012). This suggests that pitch direction sensitivity may be a marker for auditory, language, and musical abilities (Loui et al., 2008, 2011; Patel, 2008; Stewart, 2008). Interestingly, however, Mandarin-speaking amusics and controls in fact show lower pitch direction discrimination thresholds in comparison to their English-speaking counterparts, presumably because of perceptual learning of a tone language (Liu, Jiang, et al., 2012; Liu et al., 2010). However, without conscious recognition of the direction of pitch movements (Liu, Xu, et al., 2012), Mandarin-speaking amusics still demonstrate impaired melodic contour processing compared to normal controls (Jiang et al., 2010).

Furthermore, although there has been evidence suggesting that amusics were able to process subtle pitch changes and pitch direction pre-attentively in neuroimaging, ERP (event-related potentials), and pitch imitation tasks, this implicit pitch processing ability does not seem to induce normal musical functioning in amusia (Hutchins & Peretz, 2012; Hyde et al., 2011; Liu et al., 2013, 2010; Loui et al., 2008; Mignault Goulet et al., 2012; Moreau, Jolicoeur, & Peretz, 2009; Moreau, Jolicoeur, & Peretz, 2013; Peretz et al., 2009). Thus, in the current study, we trained amusics to consciously identify pitch direction by providing explicit feedback after each trial. Although focused-attention is not necessary for perceptual learning (Seitz & Watanabe, 2005), learning with feedback is much more efficient than without feedback (Herzog & Fahle, 1998). In the current training paradigm, we used visual displays of pitch contours to help amusics develop pitch direction awareness. Given the possible link between pitch processing and spatial processing in amusia (Douglas & Bilkey, 2007; although see Tillmann et al., 2010; Williamson, Cocchini, & Stewart,

2011 for different results), it will be interesting to find out whether perceptual training of complicated melodic contour patterns and their visual displays will help ameliorate musical processing deficits in amusia, and how learned patterns are encoded in auditory and visual cortical networks (Li, Piëch, & Gilbert, 2008).

Both primates and humans represent pitch direction in the right lateral Heschl's gyrus (Bendor, 2012; Bendor & Wang, 2005; Griffiths & Hall, 2012; Johnsrude, Penhune, & Zatorre, 2000; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Tramo, Cariani, Koh, Makris, & Braida, 2005). Previous studies indicate that animals such as monkeys and ferrets can be trained to discriminate pitch direction (Brosch, Selezneva, Bucks, & Scheich, 2004; Selezneva, Scheich, & Brosch, 2006; Walker, Schnupp, Hart-Schnupp, King, & Bizley, 2009). However, for humans, difficulty in pitch direction identification persists even after more than 2000 identification trials followed by visual feedback in an adaptive testing procedure for two out of three participants tested in (Semal & Demany, 2006). This may be because it takes at least 4-8 hours of training for pitch discrimination to be optimized (Micheyl, Delhommeau, Perrot, & Oxenham, 2006), and learning and memory need to be facilitated through sleep (Diekelmann, 2014). Sensitivity to pitch direction emerges from asymmetric lateral inhibition among neighboring cells in tonotopic maps (Husain, Tagamets, Fromm, Braun, & Horwitz, 2004; Ohl, Schulze, Scheich, & Freeman, 2000; Rauschecker, 1998a, 1998b; Shamma, Fleshman, Wiser, & Versnel, 1993). To our knowledge, our study is the first to systematically train a large sample of human listeners on pitch direction identification (Walker, Bizley, King, & Schnupp, 2011). Neuroimaging studies are required to explore how this behavioral improvement is linked to anatomical patterns of inhibitory connections between cells in the human auditory cortex.

Overall, our results suggest that amusics' sensitivity to pitch direction can be improved through incremental perceptual learning to a level closer to normal limits. However, pitch direction training alone may not be able to increase amusics' musical pitch perception. This stands in contrast with the transferability between pitch discrimination and speech processing (Bidelman, Gandour, & Krishnan, 2011; Bidelman, Hutka, & Moreno, 2013; Lee & Hung, 2008; Pfordresher & Brown, 2009; P. C. M. Wong, Skoe, Russo, Dees, & Kraus, 2007). Several possibilities may underlie our current results.

Firstly, previous research on humans suggests that training on pitch discrimination at certain frequencies, with different timbres, or across different durations and ears may or may not generalize to other untrained conditions (Amitay, Hawkey, & Moore, 2005; Delhommeau, Micheyl, Jouvent, & Collet, 2002; Demany, 1985; Demany & Semal, 2002; Irvine, Martin, Klimkeit, & Smith, 2000). This suggests that auditory perceptual learning may be condition-specific. As reviewed by Seitz & Watanabe (2005), task-irrelevant learning is possible only when task-irrelevant features are related to target features. For example, only when the direction of a subliminal motion is temporally-paired with the task target, can this motion be passively learned (Seitz & Watanabe, 2003). Our finding is consistent with this hypothesis, as enhanced pitch direction identification only has a subtle positive impact on musical contour processing for trained amusics, but not on musical processing as a whole. This is presumably because pitch direction processing is only a small part of musical processing (Peretz & Coltheart, 2003; Stewart, 2011). Given that pitch direction identification mainly reflects melodic contour perception, training of pitch direction may not have a direct impact on tonality (MBEA scale subtest) and pitch change (MBEA interval subtest) perception in amusia.

Furthermore, one reason that the training did not enhance amusic performance on the MBEA contour subtest to the normal level may be that the training only involved two-tone sequences, while the MBEA melodies involve longer sequences of notes (the numbers of notes in the MBEA contour subtest melodies ranged between 7 and 21, with mean = 10 and SD = 2.92). Since amusics are known to have problems with short-term memory for tone patterns (Albouy, Mattout, et al., 2013; Tillmann et al., 2009; Williamson & Stewart, 2010), it is possible that training would be more effective if amusics were adaptively trained on pitch direction tasks that involved longer tone sequences. Thus, one strategy for future training studies would be to introduce 3-tone sequences to amusics after they reach normal thresholds for two-tone sequences, then once they master those, introduce 4-tone sequences, and so on.

Alternatively, our finding that the trained amusics achieved pitch direction identification thresholds similar to the normal level, but remained within the amusic range for the MBEA pitch-based subtests suggests that pitch direction deficits may not be the sole cause for amusia, and fine-grained pitch perception may also play an important role in musical processing (Vuvan, Nunes-Silva, & Peretz, 2015). It is likely that amusia emerges from a combination of deficits, e.g., a pitch change/direction deficit, a tonal memory deficit, and a deficit with conscious access to implicit knowledge of musical patterns. That is, the melodic contour deficit may only be part of the picture. Further training studies comparing different strategies/designs are required to confirm this hypothesis.

Apart from a wide range of auditory and musical impairments, amusics also showed difficulties in learning frequencies and conditional probabilities of pitch events in tonal sequences (Loui & Schlaug, 2012; Peretz, Saffran, Schön, & Gosselin, 2012; but see Omigie & Stewart, 2011 for different results). Furthermore, although



amusic demonstrated implicit processing of melodic structure/expectation and harmonic structure in Western music, they were unable to perform as well as controls in an explicit manner (Albouy, Schulze, Caclin, & Tillmann, 2013; Jiang et al., 2016; Omigie et al., 2012; Tillmann et al., 2012). Further studies are required to use the scaffolding/incremental learning approach to train amusic on other aspects of auditory/musical processing, especially in an explicit manner. In addition, given the link between language learning and music learning (Herholz & Zatorre, 2012; Loui et al., 2011; Patel, 2011), it will be interesting to examine whether and to what extent our training paradigm in pitch direction identification can be used to facilitate language learning in second language acquisition (Chandrasekaran, Kraus, & Wong, 2012; Chandrasekaran, Sampath, & Wong, 2010), and to treat other learning disabilities such as developmental dyslexia (Besson et al., 2007; Loui et al., 2011; Ziegler et al., 2012).

Finally, it is worth noting that the current study is only an initial attempt to improve pitch direction processing in amusia through auditory training. In particular, in order to optimize learning effects in amusia, we used the same stimuli and test procedure in pre- and post-tests, which allowed direct comparisons between tasks and groups. Future studies are required to explore whether amusic are able to learn to perform cognitively more demanding tasks such as introducing roving of reference frequency in pitch direction identification (Mathias et al., 2010, 2011) and training of more complex pitch patterns in longer tonal sequences (Foxton, Brown, Chambers, & Griffiths, 2004).

## **5. Conclusion**

In summary, the current study provides the first evidence suggesting that the ability to identify pitch direction in music and speech can be improved through

616 perceptual learning in humans such as those with congenital amusia. However, the  
617 enhanced ability to identify pitch direction does not seem to have a direct beneficial  
618 effect on musical processing in amusia. Overall, these findings suggest that  
619 neurodevelopmental disabilities such as congenital amusia may be tackled through  
620 incremental learning of small components in musical processing via a scaffolding  
621 approach, which may build the base for successful learning of more complex musical  
622 systems.  
623

## References

- Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P.-E., ... Tillmann, B. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain*, *136*, 1639–1661.
- Albouy, P., Schulze, K., Caclin, A., & Tillmann, B. (2013). Does tonality boost short-term memory in congenital amusia? *Brain Research*, *1537*, 224–232.
- Allen, G. (1878). Note-Deafness. *Mind*, *3*, 157–167.
- Amitay, S., Hawkey, D. J. C., & Moore, D. R. (2005). Auditory frequency discrimination learning is affected by stimulus variability. *Perception & Psychophysics*, *67*, 691–698.
- Anderson, S., Himonides, E., Wise, K., Welch, G., & Stewart, L. (2012). Is there potential for learning in amusia? A study of the effect of singing intervention in congenital amusia. *Annals of the New York Academy of Sciences*, *1252*, 345–353.
- Ayotte, J., Peretz, I., & Hyde, K. L. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain: A Journal of Neurology*, *125*, 238–251.
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*, 379–384.
- Bendor, D. (2012). Does a pitch center exist in auditory cortex? *Journal of Neurophysiology*, *107*, 743–746.
- Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, *436*, 1161–1165.

647 Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of  
648 musical expertise and musical training on pitch processing in music and  
649 language. *Restorative Neurology and Neuroscience*, 25, 399–410.

650 Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of  
651 music and language experience on the representation of pitch in the human  
652 auditory brainstem. *Journal of Cognitive Neuroscience*, 23, 425–434.

653 Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone Language Speakers and  
654 Musicians Share Enhanced Perceptual and Cognitive Abilities for Musical  
655 Pitch: Evidence for Bidirectionality between the Domains of Language and  
656 Music. *PLoS ONE*, 8, e60676.

657 Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer.  
658 *Glott International*, 5, 341–345.

659 Brosch, M., Selezneva, E., Bucks, C., & Scheich, H. (2004). Macaque monkeys  
660 discriminate pitch relationships. *Cognition*, 91, 259–272.

661 Chandrasekaran, B., Kraus, N., & Wong, P. C. M. (2012). Human inferior colliculus  
662 activity relates to individual differences in spoken language learning. *Journal*  
663 *of Neurophysiology*, 107, 1325–1336.

664 Chandrasekaran, B., Sampath, P. D., & Wong, P. C. M. (2010). Individual variability  
665 in cue-weighting and lexical tone learning. *The Journal of the Acoustical*  
666 *Society of America*, 128, 456–465.

667 Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 edition).  
668 Hillsdale, N.J: Routledge.

669 Delhommeau, K., Micheyl, C., Jouvent, R., & Collet, L. (2002). Transfer of learning  
670 across durations and ears in auditory frequency discrimination. *Perception &*  
671 *Psychophysics*, 64, 426–436.

672 Demany, L. (1985). Perceptual learning in frequency discrimination. *The Journal of*  
673 *the Acoustical Society of America*, 78, 1118–1120.

674 Demany, L., & Semal, C. (2002). Learning to perceive pitch differences. *The Journal*  
675 *of the Acoustical Society of America*, 111, 1377–1388.

676 Diekelmann, S. (2014). Sleep for cognitive enhancement. *Frontiers in Systems*  
677 *Neuroscience*, 8. doi:10.3389/fnsys.2014.00046

678 Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial  
679 processing. *Nature Neuroscience*, 10, 915–921.

680 Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory  
681 for melodies. *Psychological Review*, 85, 341–354.

682 Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in  
683 memory for melodies. *The Journal of the Acoustical Society of America*, 49,  
684 524–531.

685 Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001).  
686 Genetic correlates of musical pitch recognition in humans. *Science (New York,*  
687 *N.Y.)*, 291, 1969–1972.

688 Duanmu, S. (2007). *The Phonology of Standard Chinese* (Second Edition). The  
689 Phonology of the World's Languages.

690 Elman, J. L. (1993). Learning and development in neural networks: the importance of  
691 starting small. *Cognition*, 48, 71–99.

692 Fancourt, A., Dick, F., & Stewart, L. (2013). Pitch-change detection and pitch-  
693 direction discrimination in children. *Psychomusicology: Music, Mind, and*  
694 *Brain*, 23, 73–81.

695 Foxton, J. M., Brown, A. C. B., Chambers, S., & Griffiths, T. D. (2004). Training  
696 improves acoustic pattern perception. *Current Biology: CB*, 14, 322–325.

697 Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004).  
698 Characterization of deficits in pitch perception underlying ‘tone deafness’.  
699 *Brain*, 127, 801–810.

700 Foxton, J. M., Nandy, R. K., & Griffiths, T. D. (2006). Rhythm deficits in ‘tone  
701 deafness’. *Brain and Cognition*, 62, 24–29.

702 Foxton, J. M., Weisz, N., Bauchet-Lecaigard, F., Delpuech, C., & Bertrand, O.  
703 (2009). The neural bases underlying pitch processing difficulties. *NeuroImage*,  
704 45, 1305–1313.

705 Francart, T., van Wieringen, A., & Wouters, J. (2008). APEX 3: a multi-purpose test  
706 platform for auditory psychophysical experiments. *Journal of Neuroscience*  
707 *Methods*, 172, 283–293.

708 Geschwind, N. (1984). The brain of a learning-disabled individual. *Annals of*  
709 *Dyslexia*, 34, 319–327.

710 Gilbert, C. D., & Li, W. (2012). Adult Visual Cortical Plasticity. *Neuron*, 75, 250–  
711 264.

712 Goldowsky, B. N., & Newport, E. L. (1993). Modeling the effects of processing  
713 limitations on the acquisition of morphology: The Less is More hypothesis. In  
714 *The Proceedings of the 11th West Coast Conference on Formal Linguistics*  
715 (pp. 124–138). Stanford, CA: CSLI.

716 Gottfried, J. A. (2008). Perceptual and neural plasticity of odor quality coding in the  
717 human brain. *Chemosensory Perception*, 1, 127–135.

718 Griffiths, T. D., & Hall, D. A. (2012). Mapping pitch representation in neural  
719 ensembles with fMRI. *The Journal of Neuroscience*, 32, 13343–13347.

720 Henry, M. J., & McAuley, J. D. (2010). On the prevalence of congenital amusia.  
721 *Music Perception*, 27, 413–418.

722 Henry, M. J., & McAuley, J. D. (2013). Failure to apply signal detection theory to the  
723 Montreal Battery of Evaluation of Amusia may misdiagnose amusia. *Music*  
724 *Perception: An Interdisciplinary Journal*, 30, 480–496.

725 Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain  
726 plasticity: behavior, function, and structure. *Neuron*, 76, 486–502.

727 Herzog, M. H., & Fahle, M. (1998). Modeling perceptual learning: difficulties and  
728 how they can be overcome. *Biological Cybernetics*, 78, 107–117.

729 Holm, S. (1979). A simple sequentially rejective multiple test procedure.  
730 *Scandinavian Journal of Statistics*, 6, 65–70.

731 Howell, D. C. (2009). *Statistical Methods for Psychology* (7 edition). Australia :  
732 Belmont, CA: Cengage Learning.

733 Husain, F. T., Tagamets, M.-A., Fromm, S. J., Braun, A. R., & Horwitz, B. (2004).  
734 Relating neuronal dynamics for auditory object processing to neuroimaging  
735 activity: a computational modeling and an fMRI study. *NeuroImage*, 21,  
736 1701–1720.

737 Hutchins, S., & Peretz, I. (2012). Amusics can imitate what they cannot discriminate.  
738 *Brain and Language*, 123, 234–239.

739 Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I.  
740 (2007). Cortical thickness in congenital amusia: when less is better than more.  
741 *The Journal of Neuroscience*, 27, 13028–13032.

742 Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological*  
743 *Science*, 15, 356–360.

744 Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an  
745 abnormal neural network for pitch processing in congenital amusia. *Cerebral*  
746 *Cortex*, 21, 292–299.

747 Idson, W. L., & Massaro, D. W. (1978). A bidimensional model of pitch in the  
748 recognition of melodies. *Perception & Psychophysics*, 24, 551–565.

749 Irvine, D. R., Martin, R. L., Klimkeit, E., & Smith, R. (2000). Specificity of  
750 perceptual learning in a frequency discrimination task. *The Journal of the*  
751 *Acoustical Society of America*, 108, 2964–2968.

752 Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic  
753 contour and speech intonation in congenital amusics with Mandarin Chinese.  
754 *Neuropsychologia*, 48, 2630–2639.

755 Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2011). Fine-grained pitch  
756 discrimination in congenital amusics with Mandarin Chinese. *Music*  
757 *Perception*, 28, 519–526.

758 Jiang, C., Lim, V. K., Wang, H., & Hamm, J. P. (2013). Difficulties with pitch  
759 discrimination influences pitch memory performance: Evidence from  
760 congenital amusia. *PLoS ONE*, 8, e79216.

761 Jiang, C., Liu, F., & Thompson, W. F. (2016). Impaired Explicit Processing of  
762 Musical Syntax and Tonality in a Group of Mandarin-Speaking Congenital  
763 Amusics. *Music Perception: An Interdisciplinary Journal*, 33, 401–413.

764 Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the  
765 right human auditory cortex for perceiving pitch direction. *Brain*, 123, 155–  
766 163.

767 Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia): frequency,  
768 development, genetics and musical background. *Annals of Human Genetics*,  
769 43, 369–382.

770 Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal  
771 functions in music. *Memory & Cognition*, 10, 243–251.



772 Lebrun, M.-A., Moreau, P., McNally-Gagnon, A., Mignault Goulet, G., & Peretz, I.  
 773 (2012). Congenital amusia in childhood: A case study. *Cortex; a Journal*  
 774 *Devoted to the Study of the Nervous System and Behavior*, 48, 683–688.  
 775 Lee, C.-Y., & Hung, T.-H. (2008). Identification of Mandarin tones by English-  
 776 speaking musicians and nonmusicians. *The Journal of the Acoustical Society*  
 777 *of America*, 124, 3235–3248.  
 778 Li, W., Piëch, V., & Gilbert, C. D. (2008). Learning to link visual contours. *Neuron*,  
 779 57, 442–451.  
 780 Lin, M.-C. (1988). Putonghua shengdiao de shengxue texing he zhijue Zhengzhao  
 781 [The acoustic characteristics and perceptual cues of tones in Standard  
 782 Chinese]. *Zhongguo Yuwen [Chinese Linguistics]*, 204, 182–193.  
 783 Liu, F., Jiang, C., Pfordresher, P. Q., Mantell, J. T., Xu, Y., Yang, Y., & Stewart, L.  
 784 (2013). Individuals with congenital amusia imitate pitches more accurately in  
 785 singing than in speaking: Implications for music and language processing.  
 786 *Attention, Perception & Psychophysics*, 75, 1783–1798.  
 787 Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The  
 788 mechanism of speech processing in congenital amusia: Evidence from  
 789 Mandarin speakers. *PloS One*, 7, e30374.  
 790 Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in  
 791 congenital amusia: Discrimination, identification and imitation. *Brain*, 133,  
 792 1682–1693.  
 793 Liu, F., Xu, Y., Patel, A. D., Francart, T., & Jiang, C. (2012). Differential recognition  
 794 of pitch patterns in discrete and gliding stimuli in congenital amusia: Evidence  
 795 from Mandarin speakers. *Brain and Cognition*, 79, 209–215.

796   Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception  
797       mismatch in tone-deafness. *Current Biology : CB*, 18, R331–R332.

798   Loui, P., Kroog, K., Zuk, J., & Schlaug, G. (2011). Relating pitch awareness to  
799       phonemic awareness in children: implications for tone-deafness and dyslexia.  
800       *Frontiers in Auditory Cognitive Neuroscience*, 2, 111.

801   Loui, P., & Schlaug, G. (2012). Impaired learning of event frequencies in tone  
802       deafness. *Annals of the New York Academy of Sciences*, 1252, 354–360.

803   Mathias, S. R., Bailey, P. J., Semal, C., & Demany, L. (2011). A note about  
804       insensitivity to pitch-change direction. *The Journal of the Acoustical Society*  
805       *of America*, 130, EL129–EL134.

806   Mathias, S. R., Micheyl, C., & Bailey, P. J. (2010). Stimulus uncertainty and  
807       insensitivity to pitch-change direction. *The Journal of the Acoustical Society*  
808       *of America*, 127, 3026–3037.

809   Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of  
810       musical and psychoacoustical training on pitch discrimination. *Hearing*  
811       *Research*, 219, 36–47.

812   Mignault Goulet, G., Moreau, P., Robitaille, N., & Peretz, I. (2012). Congenital  
813       amusia persists in the developing brain after daily music listening. *PLoS ONE*,  
814       7, e36860.

815   Moreau, P., Jolicoeur, P., & Peretz, I. (2009). Automatic brain responses to pitch  
816       changes in congenital amusia. *Annals of the New York Academy of Sciences*,  
817       1169, 191–194.

818   Moreau, P., Jolicoeur, P., & Peretz, I. (2013). Pitch discrimination without awareness  
819       in congenital amusia: evidence from event-related potentials. *Brain and*  
820       *Cognition*, 81, 337–344.

821 Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone  
 822 language: Association with lexical tone agnosia. *Brain*, 133, 2635–2642.

823 Neuhoﬀ, J. G., Knight, R., & Wayand, J. (2002). Pitch change, soniﬁcation, and  
 824 musical expertise: Which way is up? In *Proceedings of the International*  
 825 *Conference on Auditory Display* (pp. 351–356).

826 Newport, E. L. (1988). Constraints on learning and their role in language acquisition:  
 827 Studies of the acquisition of American sign language. *Language Sciences*, 10,  
 828 147–172.

829 Ohl, F. W., Schulze, H., Scheich, H., & Freeman, W. J. (2000). Spatial representation  
 830 of frequency-modulated tones in gerbil auditory cortex revealed by epidural  
 831 electrocorticography. *Journal of Physiology, Paris*, 94, 549–554.

832 Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics:  
 833 measures of effect size for some common research designs. *Psychological*  
 834 *Methods*, 8, 434–447.

835 Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in  
 836 congenital amusia. *Neuropsychologia*, 50, 1483–1493.

837 Omigie, D., & Stewart, L. (2011). Preserved statistical learning of tonal and linguistic  
 838 material in congenital amusia. *Frontiers in Psychology*, 2, 109.

839 Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford University Press, USA.

840 Patel, A. D. (2011). Why would musical training benefit the neural encoding of  
 841 speech? The OPERA hypothesis. *Frontiers in Psychology*, 2, 142.

842 Patel, A. D., Foxton, J. M., & Griﬃths, T. D. (2005). Musically tone-deaf individuals  
 843 have diﬃculty discriminating intonation contours extracted from speech.  
 844 *Brain and Cognition*, 59, 310–313.

845 Patel, A. D., Wong, M., Foxton, J. M., Lochy, A., & Peretz, I. (2008). Speech  
846 intonation perception deficits in musical tone deafness (congenital amusia).  
847 *Music Perception*, 25, 357–368.

848 Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The  
849 processing of temporal pitch and melody information in auditory cortex.  
850 *Neuron*, 36, 767–776.

851 Peretz, I. (2013). The biological foundations of music: Insights from congenital  
852 amusia. In *The Psychology of Music (Third Edition)* (pp. 551–564). Academic  
853 Press.

854 Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B.  
855 (2002). Congenital amusia: a disorder of fine-grained pitch discrimination.  
856 *Neuron*, 33, 185–191.

857 Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In  
858 tune, out of key, and unaware. *Brain*, 132, 1277–1286.

859 Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of musical disorders. The  
860 Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy*  
861 *of Sciences*, 999, 58–75.

862 Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature*  
863 *Neuroscience*, 6, 688–691.

864 Peretz, I., Cummings, S., & Dubé, M.-P. (2007). The genetics of congenital amusia  
865 (tone deafness): A family-aggregation study. *American Journal of Human*  
866 *Genetics*, 81, 582–588.

867 Peretz, I., Saffran, J., Schön, D., & Gosselin, N. (2012). Statistical learning of speech,  
868 not music, in congenital amusia. *Annals of the New York Academy of Sciences*,  
869 1252, 361–367.

870 Peron, R. M., & Allen, G. L. (1988). Attempts to train novices for beer flavor  
 871 discrimination: a matter of taste. *The Journal of General Psychology*, 115,  
 872 403–418.

873 Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of  
 874 musical pitch in tone language speakers. *Attention, Perception, &*  
 875 *Psychophysics*, 71, 1385–1398.

876 R Core Team. (2014). *R: A language and environment for statistical computing*.  
 877 Vienna, Austria: R Foundation for Statistical Computing. Retrieved from URL  
 878 <http://www.R-project.org/>

879 Rauschecker, J. P. (1998a). Cortical processing of complex sounds. *Current Opinion*  
 880 *in Neurobiology*, 8, 516–521.

881 Rauschecker, J. P. (1998b). Parallel processing in the auditory cortex of primates.  
 882 *Audiology & Neuro-Otology*, 3, 86–103.

883 Seitz, A. R., & Watanabe, T. (2003). Psychophysics: Is subliminal learning really  
 884 passive? *Nature*, 422, 36–36.

885 Seitz, A. R., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends*  
 886 *in Cognitive Sciences*, 9, 329–334.

887 Selezneva, E., Scheich, H., & Brosch, M. (2006). Dual time scales for categorical  
 888 decision making in auditory cortex. *Current Biology: CB*, 16, 2428–2433.

889 Semal, C., & Demany, L. (2006). Individual differences in the sensitivity to pitch  
 890 direction. *The Journal of the Acoustical Society of America*, 120, 3907–3915.

891 Shamma, S. A., Fleshman, J. W., Wiser, P. R., & Versnel, H. (1993). Organization of  
 892 response areas in ferret primary auditory cortex. *Journal of Neurophysiology*,  
 893 69, 367–383.

894 Stewart, L. (2008). Fractionating the musical mind: insights from congenital amusia.  
895 *Current Opinion in Neurobiology*, 18, 127–130.

896 Stewart, L. (2011). Characterizing congenital amusia. *Quarterly Journal of*  
897 *Experimental Psychology* (2006), 64, 625–638.

898 Tillmann, B., Gosselin, N., Bigand, E., & Peretz, I. (2012). Priming paradigm reveals  
899 harmonic structure processing in congenital amusia. *Cortex*, 48, 1073–1078.

900 Tillmann, B., Jolicoeur, P., Ishihara, M., Gosselin, N., Bertrand, O., Rossetti, Y., &  
901 Peretz, I. (2010). The amusic brain: lost in music, but not in space. *PloS One*,  
902 5, e10173.

903 Tillmann, B., Schulze, K., & Foxton, J. M. (2009). Congenital amusia: A short-term  
904 memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition*,  
905 71, 259–264.

906 Tramo, M. J., Cariani, P. A., Koh, C. K., Makris, N., & Braida, L. D. (2005).  
907 Neurophysiology and neuroanatomy of pitch perception: auditory cortex.  
908 *Annals of the New York Academy of Sciences*, 1060, 148–174.

909 Trehub, S. E. (2010). In the beginning: A brief history of infant music perception.  
910 *Musicae Scientiae*, 14, 71–87.

911 Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the  
912 non-modularity of pitch processing in congenital amusia. *Cortex*, 69, 186–200.

913 Walker, K. M. M., Bizley, J. K., King, A. J., & Schnupp, J. W. H. (2011). Cortical  
914 encoding of pitch: Recent results and open questions. *Hearing Research*, 271,  
915 74–87.

916 Walker, K. M. M., Schnupp, J. W. H., Hart-Schnupp, S. M. B., King, A. J., & Bizley,  
917 J. K. (2009). Pitch discrimination by ferrets for simple and complex sounds.  
918 *The Journal of the Acoustical Society of America*, 126, 1321–1335.

919 Williamson, V. J., Cocchini, G., & Stewart, L. (2011). The relationship between pitch  
920 and space in congenital amusia. *Brain and Cognition*, 76, 70–76.

921 Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia:  
922 Beyond a fine-grained pitch discrimination problem. *Memory*, 18, 657–669.

923 Wong, M., Peters, R. M., & Goldreich, D. (2013). A physical constraint on perceptual  
924 learning: tactile spatial acuity improves with training to a limit set by finger  
925 size. *The Journal of Neuroscience: The Official Journal of the Society for*  
926 *Neuroscience*, 33, 9345–9352.

927 Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L.-H., & Peretz, I.  
928 (2012). Effects of culture on musical pitch perception. *PloS One*, 7, e33424.

929 Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical  
930 experience shapes human brainstem encoding of linguistic pitch patterns.  
931 *Nature Neuroscience*, 10, 420–422.

932 Wright, B. A., & Zhang, Y. (2009). A review of the generalization of auditory  
933 learning. *Philosophical Transactions of the Royal Society of London. Series B,*  
934 *Biological Sciences*, 364, 301–311.

935 Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

936 Zendel, B. R., Lagrois, M.-É., Robitaille, N., & Peretz, I. (2015). Attending to pitch  
937 information inhibits processing of pitch information: the curious case of  
938 amusia. *The Journal of Neuroscience: The Official Journal of the Society for*  
939 *Neuroscience*, 35, 3815–3824.

940 Ziegler, J. C., Pech-Georgel, C., George, F., & Foxton, J. M. (2012). Global and local  
941 pitch perception in children with developmental dyslexia. *Brain and*  
942 *Language*, 120, 265–270.

943

944 **Author note**

945         This work was supported by a grant from the National Natural Science  
946 Foundation of China (31470972) to C.J., National Science Foundation Grant BCS-  
947 1125144, the Liu Che Woo Institute of Innovative Medicine at The Chinese  
948 University of Hong Kong, the US National Institutes of Health grant R01DC013315,  
949 the Research Grants Council of Hong Kong grants 477513 and 14117514, and the  
950 Stanley Ho Medical Development Foundation. We thank Aniruddh D. Patel for  
951 insightful comments (especially in regard to interpretation of the results) on an earlier  
952 version of the manuscript, Wen Xu, Antonio Cheung, and personnel in the  
953 Bilingualism and Language Disorders Laboratory, Shenzhen Research Institute of the  
954 Chinese University of Hong Kong for their assistance in this research.

955         Correspondence concerning this article should be addressed to either Patrick  
956 C. M. Wong, Department of Linguistics and Modern Languages, The Chinese  
957 University of Hong Kong, Room G03, Leung Kau Kui Building, Shatin, N.T., Hong  
958 Kong. E-mail: p.wong@cuhk.edu.hk, or Cunmei Jiang, Music College, Shanghai  
959 Normal University, 100 E. Guilin Road, Shanghai, 200234, China. E-mail:  
960 cunmeijiang@126.com.



Table 1. Characteristics of the amusic ( $n = 20$ ; 11 female, 9 male; 1 left-handed, 19 right-handed; 18 Mandarin speakers tested in Shanghai, 2 Cantonese speakers tested in Hong Kong) and control ( $n = 20$ ; 13 female, 7 male; 1 left-handed, 19 right-handed; 18 Mandarin speakers tested in Shanghai, 2 Cantonese speakers tested in Hong Kong) groups. The trained and untrained amusic groups each contained 9 Mandarin speakers (tested in Shanghai) and 1 Cantonese speaker (tested in Hong Kong).

Group	Age	Education	Scale	Contour	Interval	Rhythm	Meter	Memory	Pitch composite	MBEA global score
<b>Amusic</b>										
Mean	23.55	17.20	17.65	18.60	16.85	22.75	18.60	21.55	53.10	64.44
SD	1.57	1.74	3.57	2.68	3.00	3.49	3.90	3.73	5.99	5.98
<b>Trained</b>										
Mean	24.20	18.20	16.80	18.40	17.20	24.20	17.00	20.50	52.40	63.39
SD	1.69	1.55	2.94	3.27	2.90	2.90	3.94	3.72	6.45	6.11
<b>Untrained</b>										
Mean	22.90	16.20	18.50	18.80	16.50	21.30	20.20	22.60	53.80	65.50
SD	1.20	1.32	4.09	2.10	3.21	3.56	3.29	3.63	5.75	5.99
<b><i>t</i>-test (T vs. U)</b>										
<i>t</i>	1.99	3.11	-1.07	-0.33	0.51	2.00	-1.97	-1.27	-0.51	-0.78
<i>p</i>	0.06	0.01	0.30	0.75	0.61	0.06	0.06	0.22	0.61	0.45
<b>Control</b>										
Mean	23.25	17.85	28.05	28.10	27.30	27.90	26.75	28.85	83.45	92.75
SD	1.71	1.81	1.23	1.25	1.89	2.00	2.47	0.93	3.38	3.65
<b><i>t</i>-test (A vs. C)</b>										
<i>t</i>	0.58	-1.16	-12.30	-14.35	-13.18	-5.72	-7.90	-8.48	-19.73	-18.06
<i>p</i>	0.57	0.25	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001

T = trained; U = untrained; A = amusic; C = control; age and education are in years; scores on the six MBEA subtests (scale, contour, interval, rhythm, meter, and memory; Peretz et al., 2003) are in number of correct responses out of 30; the pitch composite score is the sum of the scale, contour, and interval scores; MBEA global score is the percentage of correct responses out of the total 180 trials; *t* is the statistic of the Welch two sample *t*-test (two-tailed,  $df = 18$  for trained versus untrained amusics and  $df = 38$  for amusics versus controls).

## Figure captions

**Fig. 1.** Illustration of the pitch threshold tasks. The dotted line represents the reference frequency at 131 Hz (C3), and the solid lines represent the auditory stimuli (/ma/ or piano tones). The stimuli and the inter-stimulus-interval were all 250 ms in duration.

**Fig. 2.** Mean pitch thresholds (in st, or semitones) of the 20 controls, 10 trained, and 10 untrained amusics for piano tones (A) and speech syllables (B) in pre- and post-tests. Controls are denoted by dark gray squares and solid lines, trained amusics by light gray triangles and solid lines, and untrained amusics by black dots and dashed lines. Error bars represent standard errors.

**Fig. 3.** Mean pitch thresholds (in st, or semitones) across the 10 training sessions for the 10 trained amusics. Thresholds for piano tones are represented by gray squares, and those for speech syllables are denoted by black triangles. Error bars represent standard errors.

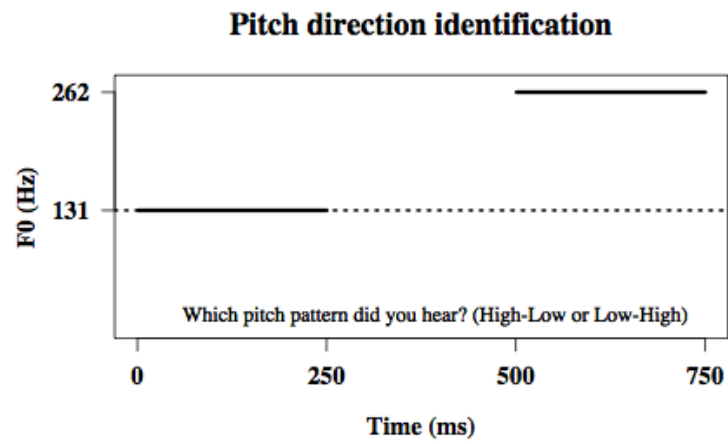
**Fig. 4.** Scatter plots of pre- versus post-test pitch thresholds (in st, or semitones) of the 10 trained and 10 untrained amusics for piano tones (A) and speech syllables (B). Untrained amusics are represented by black dots and dashed lines, and trained amusics are denoted by gray triangles and solid lines. Regression lines were based on linear regressions between pre- and post-test thresholds of trained/untrained amusics.

**Fig. 5.** Mean scores (in number of correct responses out of 30) of the 10 trained and 10 untrained amusics for MBEA scale (A), contour (B), and interval subtest (C) in

pre- and post-tests. Untrained amusics are represented by black dots and dashed lines, and trained amusics are denoted by gray triangles and solid lines. Error bars represent standard errors.

**Fig. 6.** Scatter plots of pre- versus post-test MBEA scores of the 10 trained and 10 untrained amusics for the scale (A), contour (B), and interval (C) subtests. Untrained amusics are represented by black dots and dashed lines, and trained amusics are denoted by gray triangles and solid lines. Regression lines were based on linear regressions between pre- and post-test thresholds of trained/untrained amusics. There were no significant correlations between pre- and post-test MBEA scores for either trained or untrained amusics across all three subtests, presumably due to random variations within and across participants.

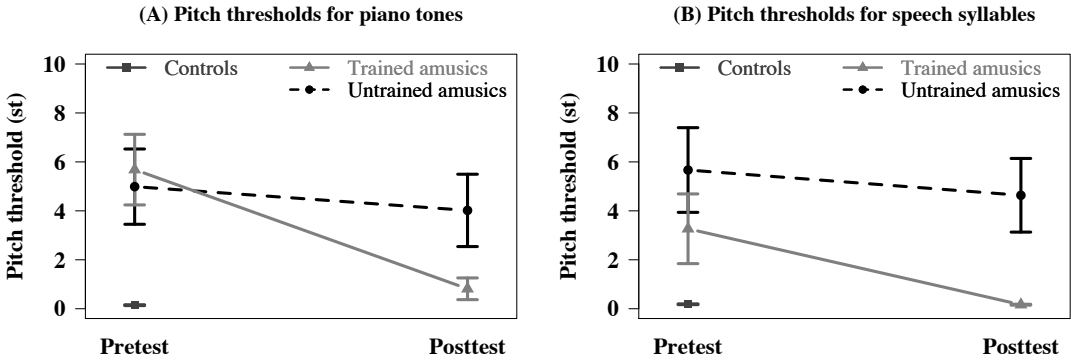
1013     Figure 1.



1014

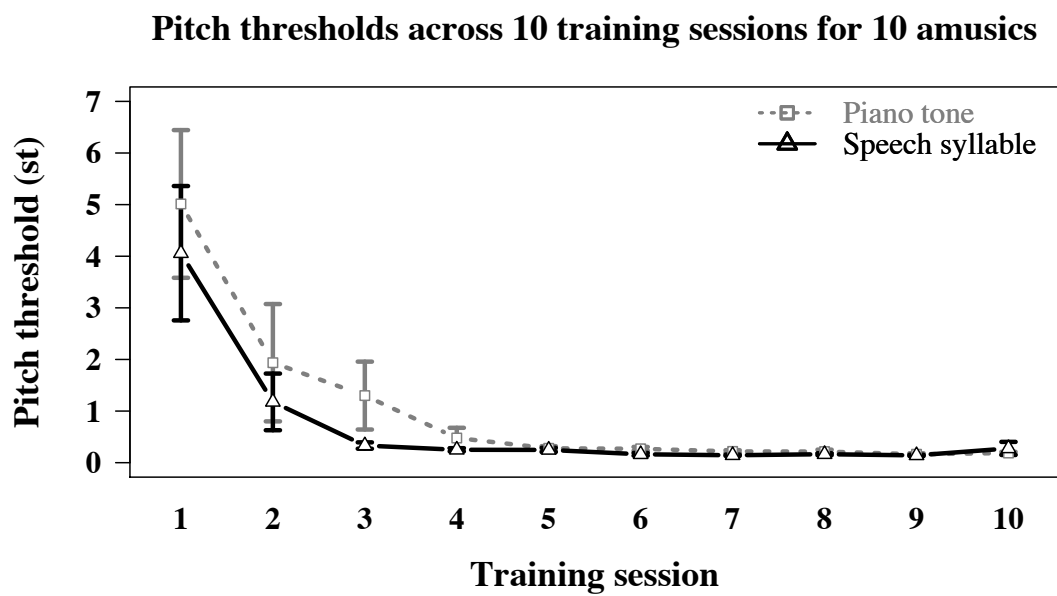
1015

1016     Figure 2.



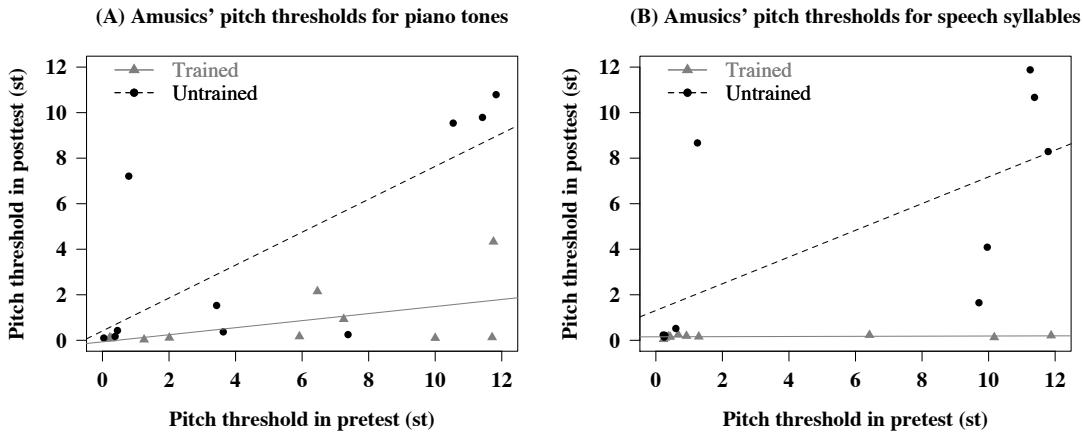
1017  
1018

1019 Figure 3.



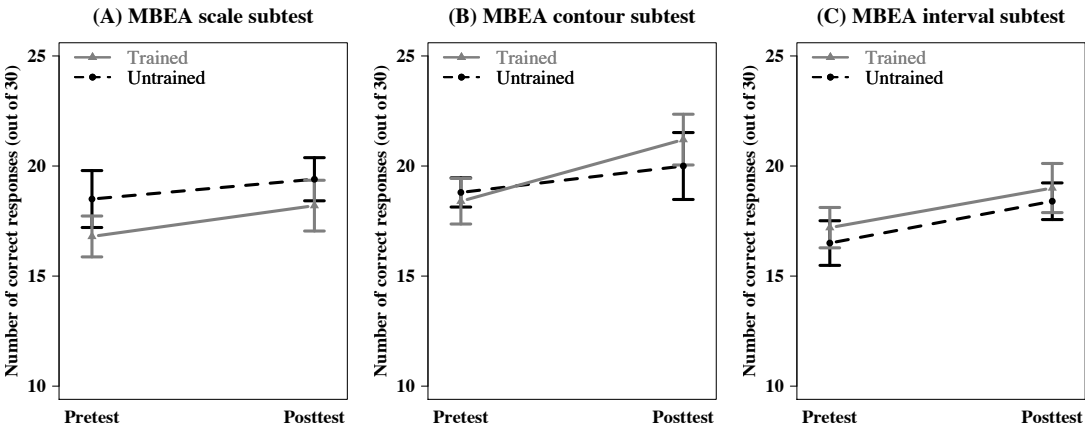
1020

1021     Figure 4.



1022

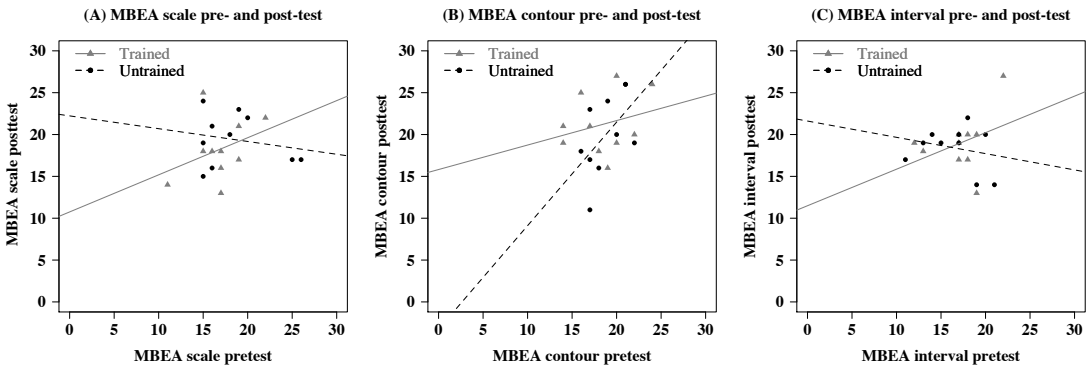
1023      Figure 5.



1024



1025      Figure 6.



1026

1027

## Supplementary materials

**Supplementary Table 1.** The linear mixed-effects model on log-transformed thresholds of the two amusic groups, with training (trained versus untrained) as the between-subjects factor, education as a covariate, stimulus type (speech syllable versus piano tone) and test (pretest versus posttest) as within-subjects factors, and participants (trained and untrained amusics) as random effects. Significant effects are in boldface.

Fixed effects	numDF	denDF	F-value	<i>p</i> -value
Intercept	1	48	3.3521	0.0733
Stimulus type	1	48	2.2065	0.1440
<b>Training</b>	<b>1</b>	<b>16</b>	<b>16.4564</b>	<b>0.0009</b>
<b>Test</b>	<b>1</b>	<b>48</b>	<b>30.4232</b>	<b>&lt;.0001</b>
Education	1	16	2.8509	0.1107
<b>Stimulus type : Training</b>	<b>1</b>	<b>48</b>	<b>7.1732</b>	<b>0.0101</b>
Stimulus type : Test	1	48	1.1501	0.2889
<b>Training : Test</b>	<b>1</b>	<b>48</b>	<b>18.4963</b>	<b>0.0001</b>
Stimulus type : Education	1	48	0.4483	0.5064
Training : Education	1	16	0.2415	0.6298
Test : Education	1	48	1.6483	0.2053
Stimulus type : Training : Test	1	48	0.0745	0.7860
Stimulus type : Training : Education	1	48	0.9521	0.3341
Stimulus type : Test : Education	1	48	2.2959	0.1363
Training : Test : Education	1	48	2.3876	0.1289
Stimulus type : Training : Test : Education	1	48	1.0788	0.3042

**Supplementary Table 2.** The linear mixed-effects model on posttest threshold (log-transformed), with training (trained versus untrained) and stimulus type (piano tone versus speech syllable) as fixed effects, pretest threshold (log-transformed) and education as covariates, and participants (trained and untrained amusics) as random effects. Significant effects are in boldface.

Fixed effects	numDF	denDF	F-value	p-value
Intercept	1	16	37.1468	<.0001
Stimulus type	1	8	0.0746	0.7917
<b>Training</b>	<b>1</b>	<b>16</b>	<b>135.5650</b>	<b>&lt;.0001</b>
<b>Pretest threshold</b>	<b>1</b>	<b>8</b>	<b>54.8023</b>	<b>0.0001</b>
<b>Education</b>	<b>1</b>	<b>16</b>	<b>18.3555</b>	<b>0.0006</b>
Stimulus type : Training	1	8	0.7948	0.3987
Stimulus type : Pretest threshold	1	8	0.7793	0.4031
<b>Training : Pretest threshold</b>	<b>1</b>	<b>8</b>	<b>26.8712</b>	<b>0.0008</b>
Stimulus type : Education	1	8	2.4113	0.1591
Training : Education	1	16	3.1772	0.0937
Pretest threshold : Education	1	8	3.5204	0.0975
<b>Stimulus type : Training : Pretest threshold</b>	<b>1</b>	<b>8</b>	<b>6.5517</b>	<b>0.0337</b>
Stimulus type : Training : Education	1	8	1.3587	0.2773
Stimulus type : Pretest threshold : Education	1	8	0.0000	0.9974
Training : Pretest threshold : Education	1	8	2.2273	0.1739
Stimulus type : Training : Pretest threshold : Education	1	8	4.1485	0.0761

**Supplementary Table 3.** The linear mixed-effects model on posttest MBEA score, with training (trained versus untrained) and task (scale, contour, and interval) as fixed effects, pretest score and education as covariates, and participants (trained and untrained amusics) as random effects. Significant effects are in boldface.

Fixed effects	numDF	denDF	F-value	<i>p</i> -value
Intercept	1	20	1292.8169	<.0001
Task	2	20	2.0625	0.1533
Training	1	16	0.3723	0.5503
Pretest score	1	20	0.0480	0.8289
<b>Education</b>	<b>1</b>	<b>16</b>	<b>7.2573</b>	<b>0.0160</b>
Task : Training	2	20	1.1059	0.3503
Task : Pretest score	2	20	1.9286	0.1714
Training : Pretest score	1	20	2.0119	0.1715
Task : Education	2	20	0.5536	0.5834
Training : Education	1	16	0.5507	0.4688
<b>Pretest score : Education</b>	<b>1</b>	<b>20</b>	<b>5.2751</b>	<b>0.0326</b>
Task : Training : Pretest score	2	20	2.0141	0.1596
Task : Training : Education	2	20	0.1641	0.8498
Task : Pretest score : Education	2	20	2.5664	0.1018
Training : Pretest score : Education	1	20	0.0269	0.8714
Task : Training : Pretest score : Education	2	20	0.5467	0.5873